

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problems Mailbox.**

FIGURE 1

TTCATCATG	GAATATTCTA	AACATACAGA	AAAATCACAG	AAAATAATA	ACAACCACTC	ATTTATCTTC	-1101
TCCCCAACCC	CATGTAAATA	ATATTAAAAT	ATTGTGTTAA	ATCTTAAATT	TAACACATGC	TAAGGGTTC	-1031
TGGTGGATG	TGGTGGCTCA	CGCCTGTAAAT	CCCAGTACTT	TGGGAGGAGG	AGGTGGGAGG	ATTGCTTGAG	-961
TCCAGGAGCT	CGAGACCCAG	ATGGCAACA	TAGTCCGATC	TGCTCTCTAC	AAAAAAACAA	AAAATAGCT	-891
GGGCATGGG	GTGTCATCA	GTAACTCCAG	TGACTGGAG	GCTGAGCTGG	GAGAATGCT	TGAGTCTGGG	-821
AATTGAGGC	TGGACTGAGC	CCTGTATCG	CCACTGCATT	CCAGCATGGG	CCACATAGCA	AAACTCTCA	-751
AAAAAAAAAA	AAGTTTCTC	TCTGCCAC	CATAGAACAC	CACTCTCTG	ATTTATATCT	TCGTAGATGA	-681
AP1							
ATTTGCCCA	TTCTCTTGT	TATGAAAGGA	ACCAGACATT	AGGCATTCTG	GTGCTGTT	TCTTCACTT	-611
AAGATAAAAT	TGAGTTAAC	TGTATGTTG	TACAGAACTG	CAGTTGTTTC	TTTGTATTT	ATTGAAAGA	-541
CAGGGTCTGG	CTATGTTGCC	TAGGCTGGTC	TCGAATGTT	GGCTCAAGC	AATCCACCTG	CCAAGCTCTG	-471
GGACCACACG	CATGAGCCAT	GGCATCTGAT	CGKTAGTTTG	ATCTTATTT	TTGCTGAGTA	GTAGCCCCATG	-401
AP1							
GCATGACTTT	ATTATTTTGG	GTGTCATTC	TCCTCTGGAG	GGGCTCTGCT	TTTGTAAACC	ACACCCCTGGC	-331
Ets							
CTAGCTCCC	TTCTCCCCCTGC	CTCTCTGCAG	GCTCACATCC	ACATGCCAAG	ACCTCTGCAG	CCATTCTGCT	-261
Ets							
TCCCTGCTT	CCACTCCCTG	GGGACCTCAG	AGAGCTACGG	GGCTCCCTGG	GTACCAACTG	GCTCTGAGG	-191
Sp1/Sp3							
<u>CCTGGGGAG</u>	<u>GGTGGTCTTC</u>	<u>TGGGAGAAGG</u>	<u>AAGCCAGGTC</u>	<u>CCTGCAGGTT</u>	<u>GTGGAGGGGG</u>	<u>ACAGAATGAG</u>	-121
Sp1/Sp3							
GGTTTTTCCC	CAGGATTTG	TTGGCCCTG	CCCCCACTTC	TGTTCCATAA	TAAACCACGC	<u>CCCTCTTACCC</u>	-51
Sp1/Sp3							
CACTGTGCC	CTCTTCTGC	TGTGTCGAGG	CCCTGAATCA	TTATTTAAC	TACCCCCCTGG	GAGGGTGTAGC	20
Ets							
ACCTCTGTG	CTCTGTCCCC	<u>ACACCTCCAC</u>	TTCCCTCAA	CGCGCTGCTC	ACGGATGACC	TTGGCACTG	90
M T F G T							
TCGTTCTTCT	GAGTCgtaa	tggggccagg	gtgtctggga	gaagcttgg	ggagttctga	ggggactcca	160
V L L L S							
tctgggaggg	caggctgggg	gttgtggtc	ggctccaacc	actctttaga	ggagctgagg	caggggagt	230
cttcatgtc	taatccccgc	actttttggag	ggcgagatgg	gctgtatccc	ttaggttcagg	agtttgcac	300
ggttgcgact	aactacactc	ttccatgtc	ttccaaaaatt	ataaaaaatt	agccccggat	gttgtgcgt	370
ctgggtgtgg	tgacagaggt	ctccccggcc	ttccccggga	gttggggc	ttctcccaact	catggagtcc	440
ggaaaagggt	gacatcacct	ccctggcct	nnnnntcccc	ccaaagttctg	actgcacgta	ggggagaggc	510
INTRON 1							
ccccctgtca	aaactgcac	agagtcacat	tcacgtgc	tcaaaaatca	ggtttggctg	ggtgtgggt	580
ctcatgtca	taatccccgc	actttttggag	ggcgagatgg	gctgtatccc	ttaggttcagg	agtttgcac	650
cagectggcc	aacatgtca	aaacctccat	ttccaaaaatt	ataaaaaatt	agccccggat	gttgtgcgt	720
acttgcata	ccagctactt	gggaagctga	ggcaagagaga	tcgttgcac	ccaggagacg	gaagttgcac	790
tgagctgaga	tcgttgcgtt	gcactccgc	cteagcaaca	gagcgagact	ccatctcaa	aaaaaaaaaaa	860
aaaaaaaagaa	aaaaaaaagaa	aagaggctgg	gaggctctag	ggattggggc	ttttttact	ccccctcc	930
ccgccccacca	aatattccctc	<u>agTCTCTGCT</u>	<u>TCTPATCAG</u>	<u>GATTCAACCT</u>	<u>GGATGTGGAG</u>	<u>GACCCCTACCA</u>	1000
V L A S Y H G F N L D V E Z P T							
TCTTCCAGGA	GGATGCAGGC	GGCTTITGGC	AGACGGCTGGT	GCATTCGGT	GGAPCTCGGt	aggccccact	1070
I F Q E D A G G F G Q S V V Q F G G S R							
INTRON 2 (3019 bp)							
cccccaagtg	ccccctgtc	ccacccctcc	tgtggctgca	gtgacatggc	catggttgt	tctccaa <u>ACT</u>	4080
L							
CGTGGTGGGA	GCACCCCTGG	AGGTGGTGGC	GGCCACCCAG	ACGGGACGGC	TGTTATGACTG	CGCAGCTGGC	4150
V V G A P L E V V A A N Q T G R L Y D C A A A							
ACGGGATGT	GGCAGCCCAT	CCCGCTGAC	<u>Ag</u> tgagtgcac	cacctggaa	ttggggccct	caacccctct	4220
T G M C Q P I P L H							
INTRON 3							
ggaccccaact	gtgccccccgc	tttagttcca	gtccagacat	tcceccggaaa	tgagtgtgt	ctgtgagtga	4290
gaccccgct	gtctgcctt	geagTCCGCC	CTGAGGCCGT	GAACATGTCC	TTGGCCCTGA	CCCTGGCAGC	4360
I R P E A V N M S L G L T L A A							
CTCCACCAAC	GGCTCCCGGC	TCCTGTgttag	tgagtgtctt	ggggccacggg	gggggtgggt	ggggccggggg	4430
S T N G S R L L							
INTRON 4							
gtgtttgtgg	ggaggagggct	ggggctggga	gtgaaggagg	agggactct	ggctcacagg		4500
cttctgtc	<u>ca</u> gGCCTGTG	GGCCACCCCT	GCACAGACTC	TGTGGGAGA	ACTCATACTC	AAAGGGTTC	4570
A C G P T L H R V C G E N S Y S K G S							
TGCCCTCTGC	TGGCTCCCGG	CTGGGAGACT	ATCCAGACAG	TCCCGGACGC	CACGCCAGt	aggteectgg	4640
C L L L G S R W E I I Q T V P D A T P							
INTRON 5 (4267 bp)							
caggagctgc	aggagggggt	tggggccccc	cagtgcacat	cogatctcc	ccccatcccc	cacagAGTGT	8840
B C							
CCACATCAAG	AGATGCACAT	CGTCTCTG	ATTGACGCT	CTGGAAGCAT	TGACCAAAAT	GACTTTAAC	8910
P H Q E M D I V F L I D G S G S I D Q N D F N							
AGATGAAGGG	CTTGTGCAA	GCTGTATGG	GGCAAGTTGA	GGGCACTGAC	ACCCTGgtga	agactggca	8980
Q M K G F V Q A V M G Q F E G T D T L							
INTRON 6 (1255 bp)							
aacaatagta	acaggcactg	agccctgggc	cctcccccact	ggccctttgca	<u>GTGTCACTG</u>	<u>ATGCAGTACT</u>	10240
F A L M Q Y							
CAAACCTCT	GAAGATCCAC	TTCACTTCA	CCCAATTCCG	GIACCAAGCCG	ACCCAGCAGA	GCCTGGTGG	10310
S N L L K I H F T P T Q F R T S P S Q Q S L V D							
TECCATCTGC	CAACTGAAG	GCCTGAGCTT	CACGGCCACG	GGCATCTCTGA	CASTGGTgtga	aagcaacccc	10380
P I V Q L K G L T F T A T G I L T V							
gaccggca	INTRON 7						

FIGURE 2

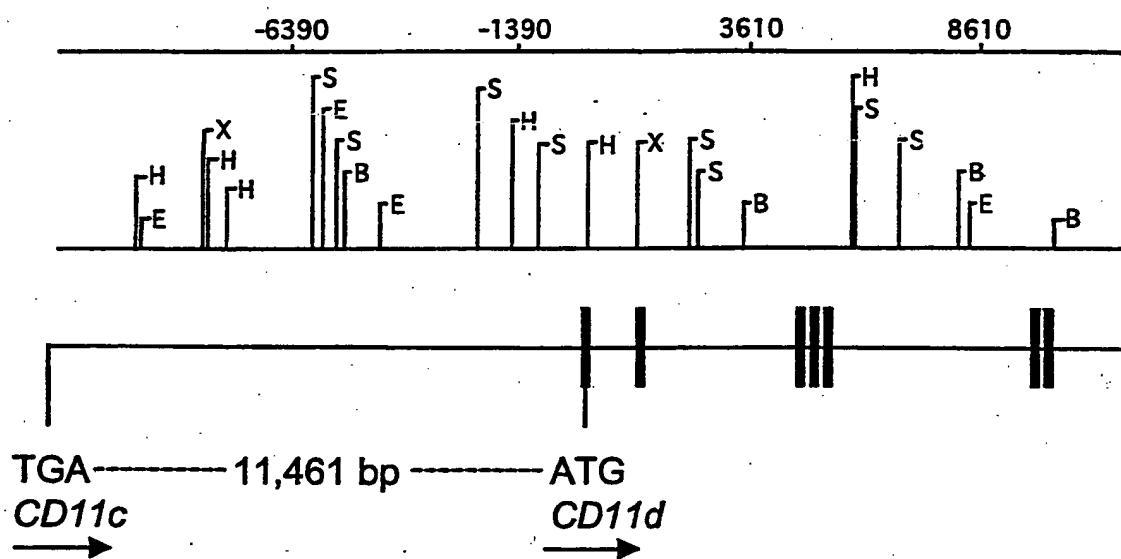


FIGURE 3

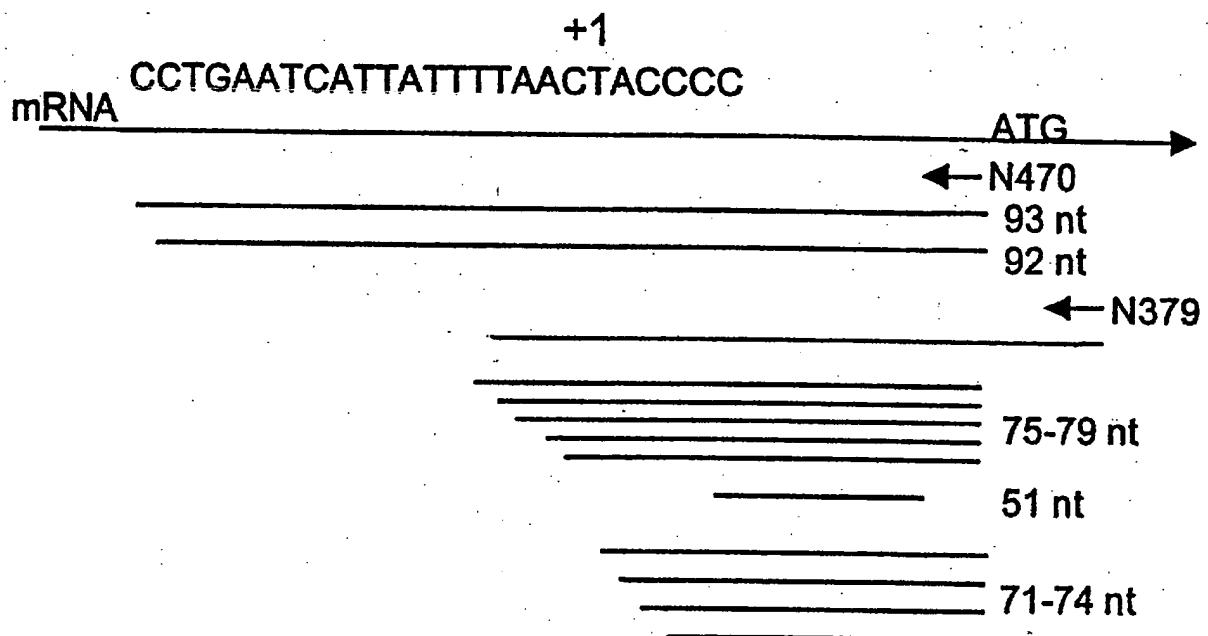


FIGURE 4

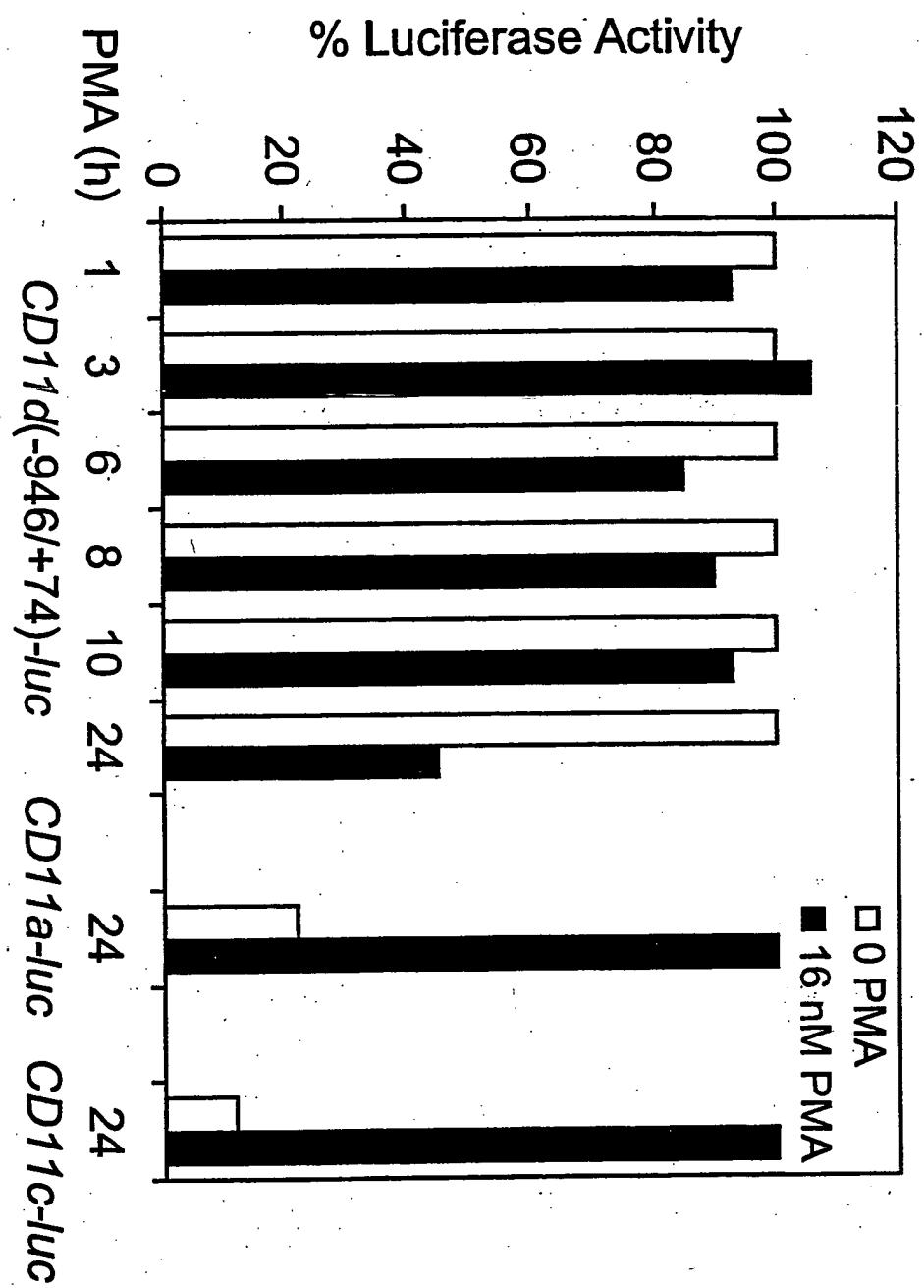


FIGURE 5

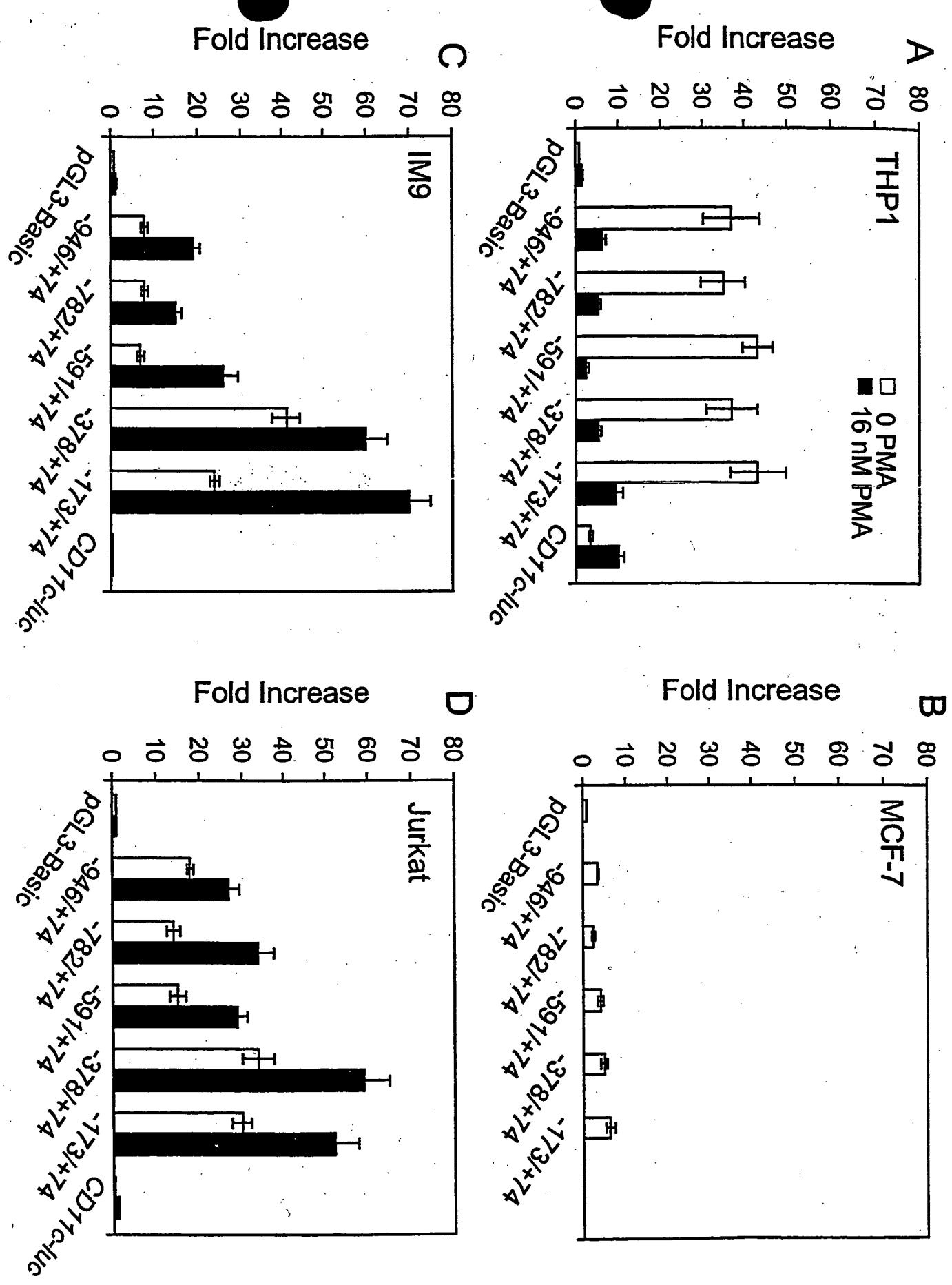


FIGURE 6

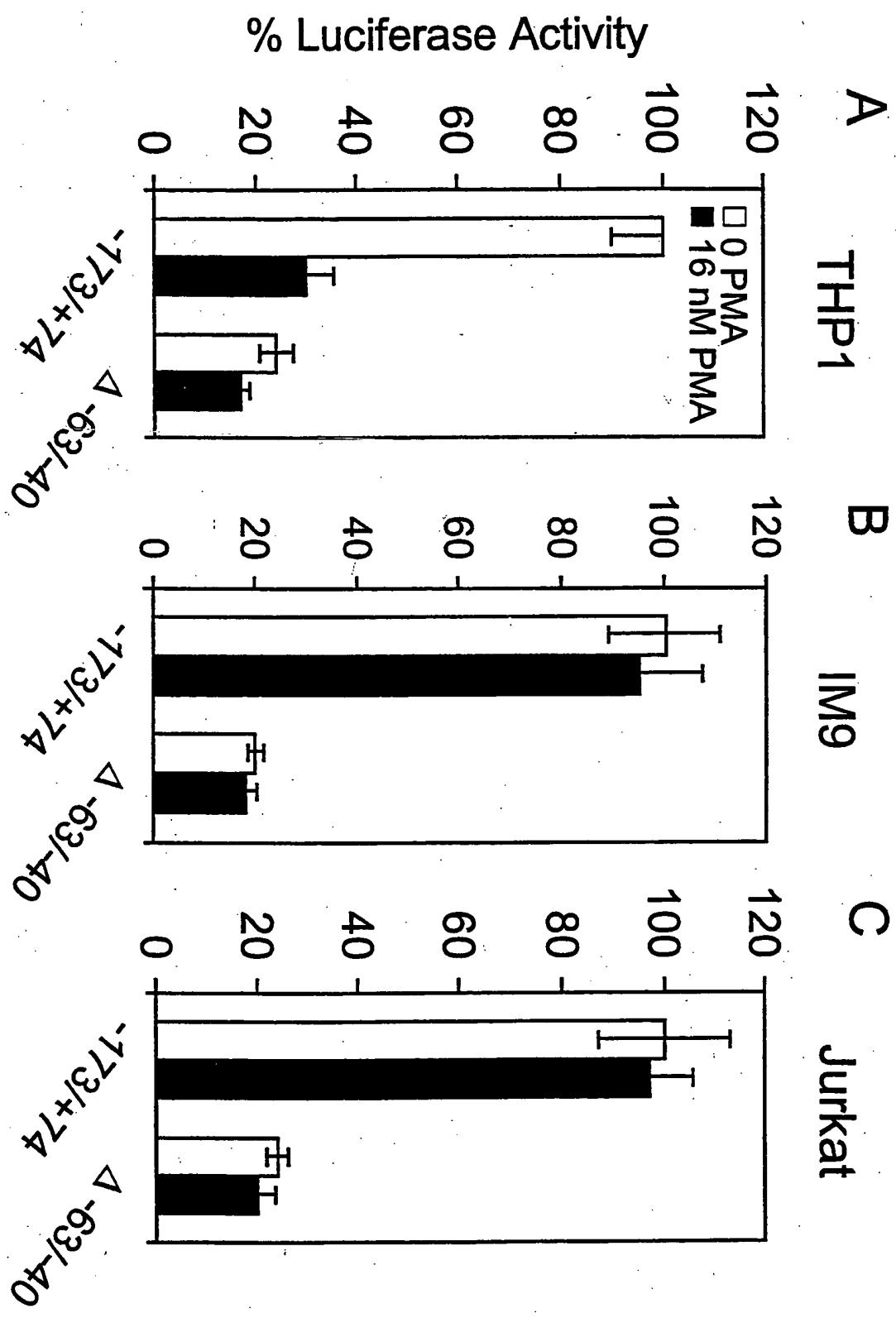


FIGURE 7

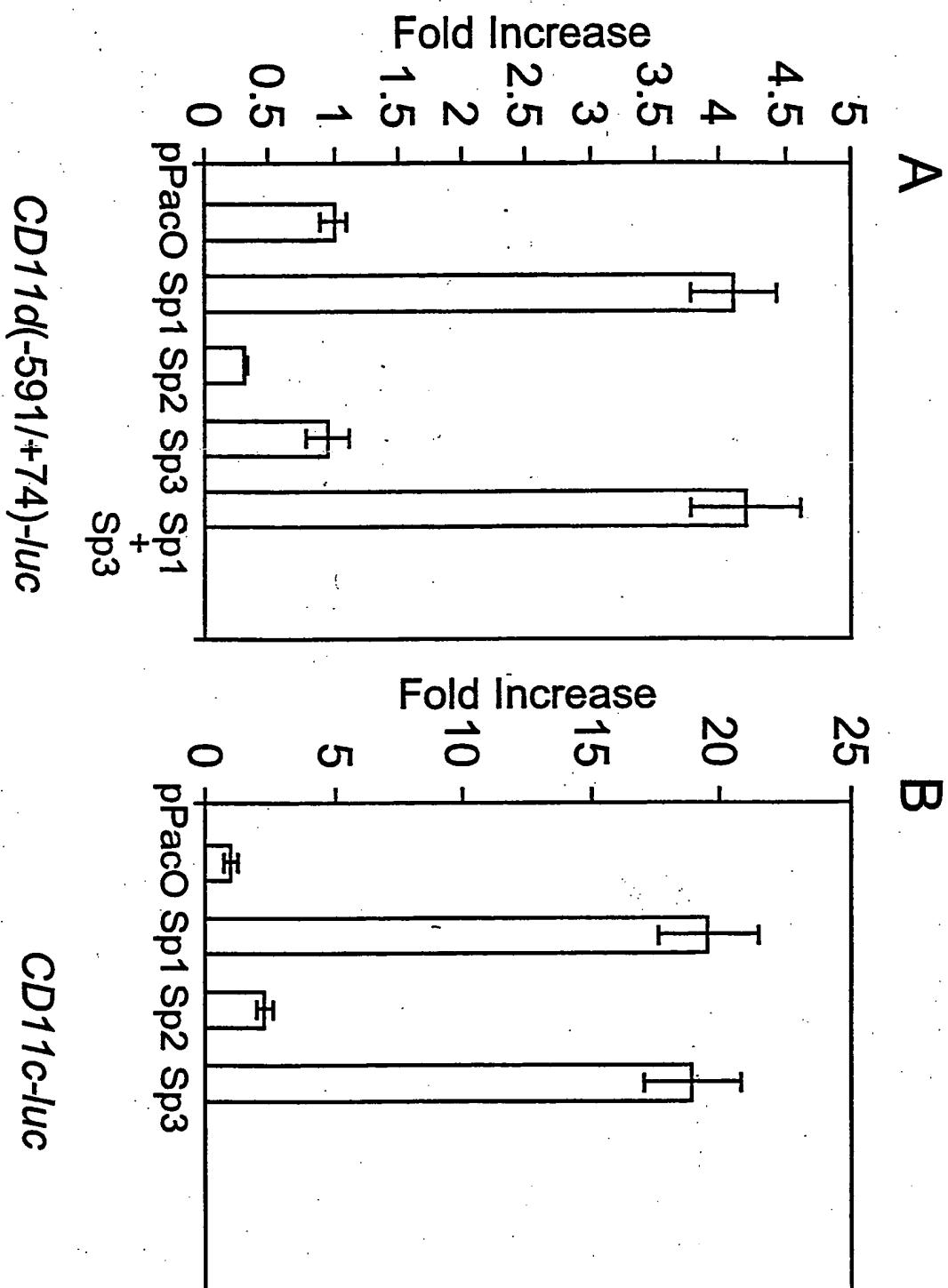


FIGURE 8

1 2 3 4 5 6 7 8 9 10
-76 -66 -54
| | |
CCATAATTAAACCACGCCCCCTCCT
•○○○○○○○○○○
-53 -38 -31
| | |
ACCCACTGTGCCCTCTTCCTGC
○○○○○○○○○○

FIGURE 9

Sequence Range: -11390 to 10387

TRANSLATIONAL
STOP
codon
FOR
CDIIc

-11321
TGATCCCTCT TTGCCTTGG A CTTCTCTCC CGCGATTTC CCCACTTACT TACCCCTCACC TGTCAGGCTG
-11251
ACGGGGAGGA ACCACTGCAC CACCGAGAGA GGCTGGGATG GCCCTGCTTC CTGTCTTGG GAGAAAACGT
-11181
CTTGCTTGGG AAGGGGCCCT TGTCTTGCA AGGTTCCAAC TGGAAACCCCT TAGGACAGGG TCCCTGCTGT
-11111
GTTCCCCAAA AGGACTTGAC TTCGAATTTC TACCTAGAAA TACATGGACA ATACCCCCAG GCCTCAGTCT
-11041
CCCTTCTCCC ATGAGGCACG AATGATCTT CTTTCCCTTC CTTTTTTTTT TTTTTCTTTT CTTTTTTTTT
-10971
TTTTTGAGA CGGAGTCTCG CTCTGTCACC CAGGCTGGAG TCGAATGGCG TGATCTCGGC TCGCTGCAAC
-10901
CTCCGCCTCC CGGGTTCAAG TAATTCTGCT GTCTCAGCCT CCTGCGTAGC TGGGACTACA GGCACACGGC
-10831
ACCTCGCCCG GCCCGATCTT TCTAAAATAC AGTTCTGAAT ATGCTGCTCA TCCCCACCTG TCTTCAACAG
-10761
CTCCCCATTA CCCTCAGGAC AATGTCTGAA CTCTCCAGCT TCGCGTGAGA AGTCCCTTC CATCCCAGAG
-10691
GGTGGGCTTC AGGGCGCACA GCATGAGAGC CTCTGTGCC CCATCACCCCT CGTTCCAGT GAATTAGTGT
-10621
CATGTCAGCA TCAGCTCAGG GCTTCATCGT GGGGCTCTCA GTTCCGATTG CCCAGGCTGA ATTGGGAGTG
-10551
AGATGCCTGC ATGCTGGTT CTGCACAGCT GGCCTCCCGC GGTGGGTCA ACATTGCTGG CCTGGAAGGG
-10481
AGGAGCGCCC TCTAGGGAGG GACATGGCCC CGGTGCGGCT GCAGCTCACC AGCCCCAGGG GCAGAACAGA
-10411
CCCAACCACT TCCTATTTTG TGAGGCTATG AATATAGTAC CTGAAAAAAAT GCCAAGCACT AGATTATTT
-10341
TTTAAAAAGC GTACTTTAAA TGTTTGTGTT AATACACATT AAAACATGCA CAAAAAGATG CATCTACCGC
-10271
TCTTGGAAA TATGTCAAAG GGTCTAAAAA TAAAAAAGCC TTCTGTGGAT ATGAGTCCTG AAGGATGACA
-10201
CCCATGGGGT CCCTTACCA CGGTGGACCC TGGCCAGCAC TGAGGCCTGG GGCCAGGACA AGAAGTTAAC
-10131
CAGAGTAGGG TTGTGAATAT CCCTCTCTTG GAAGTAACCT GACCTCTTAA TCTGCTCACT CCACTCTCAC
-10061

GGCTGGTGC C GATGGTAAGC TGGTGGAGCT GTCGGGTGGA GGGGGCATAG AATAGAGAAG GGACAACCTC
-9991
CAGTGGCTAC TTTTCCACCT GGAAAGGTCT CTGGAGTGAC CAATACTCAC AAGCGTTCC TACAAGTCCT
-9921
AGGATGTGTT GAAGGGCACA CTGTCTGCAT ATAGTGAGTG ATTGAAGAAC ATGTTGGGT CCCACATTGA
-9851
GAGCTGCTGC CCACAATAAG GTCATTCTTG CTATTATGCC ACCATCCTGG CATAAAAGTTC ATCATGGTGC
-9781
TTGGCACTGA GCTGGGGCC TCACAGGACA AGCCATTCC GACCTCGGAG TGACGCCACT GCAGCTATCA
-9711
CCAGCAAGGG ACCCGGGCCG TGTGGATGTT TCAATTAGAA AACAGAAGG GAGGCAGTTG AGTGATTGAA
-9641
AGGGAAGATG GAAAGTGGCC CTTTACCTCC AGCCAAAAAT GTCTGTCCTA TACATCAGCA GAGGCTCCAA
-9571
AATCCCTGTG GATTTGAAG CTTTGAGTC CCCAGGATGA CTAATTATTA TGCAGTTCC TCAGAAAGGG
-9501
AATCAGAAGA TAAGGCTTG TAAGAATTCA GCCCTAATGG CTGGGCACAG TGGCTCATGC CTGTAATCCC
-9431
AGCACTTGG GAGGCCGAGG CAGGAGGATT GTTTGTGCTC AGAAATTGA GACCACCTG GGTAATATAT
-9361
TGAAACCTTG TGTCTACAAA AAAATTAAA AATTACCCAG GCATGGTGGC ATGTGCTTGT AGTCCCAGCT
-9291
ACTTGGTAGG CTGAAGCAGG AGGATCACTT GAGCCTGGGA GGTTGAGGAT ACAGTGAGCT GTGATTTGGA
-9221
CCACCACACT CCAGCCTGGG CAACAGAGAA AGATCATGTC TCAGAAAAAA AAAAAAAAT TGACCCCTAGA
-9151
GTGGTGTTC TCAAAATGTG TTCCACGAAC CACTGGTGGT CAATGATGGT CTTCTAAGTG GAAGGTTTA
-9081
GAGAAAAAGA GCAAGAAACC CATAACATCTC AAACATTGA AACTAGTGAT TTGCACAGAA ATAGTGTGT
-9011
GGCCTTAATA ATTGTGTGGC ACACGGACTC CAGGGACTAC AGTGGGTTCT TGTCTAAATT CAGGCAACAA
-8941
GTTGTTATTT TCTATTTTAT TTTATTATTA TTATTTTTG AGATAGTCTC ACTTTGTCTC CCGGGCTGGA
-8871
GTGTAGTGGC ACGATCTCGG CTCAACGCAA CCTCTGTCTC CTGGGTTCAA GTGATGCCTC TGCCTCAGCC
-8801
TCCCAAGTAG CTGGGAGTAC AGGGGCGTAC CACCATGCC ATTATTTATT ATTATTTTT GAGACAGAGT
-8731

CTCGCTCTGT CACCCAGGCT GGAGTGCAGT GGCATGATCT TGGCTCACTG CAACCTCCGC CTCCCAGGTT
-8661
CAAGTTCAAG CGATTCTCCT GCCTCAGCCT CTGGAGTAGC TGGGATTACA GGCAGGCACC ACCATTCCA
-8591
GCTAATTTT GTATTTTAG TATAGATGGG GTTCACCAT GTTGACTAGG CTGGCTCGA ACTCCTGACC
-8521
TCATGATCCG CCCTCCTCGG CCTCCGAAAG TGCTGGATT AGAGGTATCA GCCACTGTAC TTGGCCGACA
-8451
AGGTGTTATT TTCTGATATT CTTCCTTGT GTGTTATTGT GTACATTGT TACATTGCA TTTTCAGGTT
-8381
TGGCTATTGT GTGCATTAG ATCCCCGAAT CACAAAATGG ATCAATGGCT CAAAAGCATG GAAGTTGTGA
-8311
TTAAAAACTA ATCTAATTGC TACAATTTAC AATAATGTCA TCAAAGTCAA TATTGACTTT TAAATATTGA
-8241
GCCCACTGCA CGTATAGTAT AGACATGCAT ACCGGAATAA GTGATTGTGA GCCAAAACCC GAAAATATCT
-8171
AGAAGGTATT ATACTCCCTG ACAGGTAGGT TGTATTGGTT CTGACATGTA TTTGTCCCTA GTGTGCTGCC
-8101
CATTCTGAAA CTTTATCAAA CAGTCGCATG AACCTCTGAA AGCTTTGTG TTATTTCTT ATTTATTTAT
-8031
TTATTGAGAT GGAGTCTTGC TCTGTCGCC AGGCTGGAGT GCAGTGGCAT CATCTGGCT CACTGCAACC
-7961
TTTGCCTCCT GGGTTCAAGT GATTCTCCTG CCTCAGCCTC TTGAGTAGCT GGGATTACAG GCGCGCACCA
-7891
CCACGCCAG CTAATTTTG TATTTTAGT ATAGACGGGG GTTTCACCAT GTGGTCAGG CTGGTCTCGA
-7821
ACCCCTGACC TCATGATCTG CCTGCCTCAG GTAAAGCAAT AGAGATTCTT AGAACAACTG CTACATGTAG
-7751
CTTCCCTATT CAAAAGTGAT TAGTGTGTC ACCGAATACA GAGGAGACAG CAAAACCACA GTGACATAAA
~~7610~~
TCAAAGGTGC TTTTAAAGT AGCAAAAGTA GGTACAAGTC ACATAATTTC CAAGAAGCTT GTAGAAATGG
-7611
CAGTAGAGTT CATACTGCT ATTGAAAGGT TGCTTTGGC TGCAAATAAT AGAAAAAAAC AAAAGCATGT
-7541
AAGAGCAGAC AGAAGACCTT TACTCTGCAA GAGGTCAGG TGCAGGTTAG TGTTAATGC AGAGTCTCAG
-7471
CATTGACAGA TTCTTCTGA TCTTCCAATT GATCGTCCTT GCGGGGGCGG TTTAGTTCTT TCCCACGTGAC
-7401

TAGGATTGGG TCAAATTCCA TCCCCTTGGT TGCATGCAGT GCTGAGAAGG TGAGCATGTG CTTTCACAG

-7331
GCTTAATAAA AAGAGGTAGC TCCAGCCAGG TGCAGTGACT CATGCCATA ATCTCAGCAC TTTGGGAGGC

-7261
AGAGGTGGGT AGGTACACCTG AGGTCAAGGAG ATTGAGAACC AGCCTGACCA ACATGGCAA ACTCTGTCTC

-7191
TACCGAAAAT ACAAAAATTA GCTGGGCATG GTGGCAGGTG CCTGTAATCC CAGCTACTTG GGAGGCTGAG

-7121
GCAGGAGAAT CGCTTGAACC TGAGAGGTGG AAGTTACAGT GAGCTGAGGT CATGCCACTT GCACTCCAGC

-7051
CTGGGGACA GAGTAGAACT CTGTCTCAAC AAAAAAAA AAAAAAAGAG AAAAAAAAG GAGGGTAGCT

-6981
CCACCAGCCA GGAAGGTGGC AGCGCTGGTG GCTGTTGGAT AGGCTACCTA CAGTGTCTGG CAAATACTAT

-6911
GCTGAAGAC TATGCTGTGA GCAAGATTCC TTTGTGAAGG AACAGCTTGG ACATTGTGTA TGTCAGAGGT

-6841
ATACAGCAGA ATAGCAGTGA CTAACGCTTG TGTGGAGAG CAAGCATGTC ACCTCATACT TGGAATAACT

-6771
CACTGCCATA CAAAGTCTGA ATCAGCTTTC GTCTTGTGC AACACATGTA TGTGGGAGCT TTTCAGCTGC

-6701
TGAAACCTCT AGTGACAGAA AAGGAGGTTT TGTTGTTCAT TTGTAATTAA TGTTAACCTT ATGAGTGGTG

-6631
GGAGAGATAG TGAGGTAGGA GATCAGCAGG ACCTGTTTC TGGTCACAAC CCAGCTAATC AGAGCATGAT

-6561
CTGGTCAAGA TGGGATGCAC TAAAAAAACA GCCCAAACCA GCAGATGGCC AGGAAAGCAA ACTCTCATTA

-6491
CCCTCGCCAC TTATTAGCAT AAAGACACTC CCACCGGTGC CATGACAGTT TACAAATGCC ATGGAAACAC

-6421
ACCATAGCAA CGGTCAGCAA GTTACCTCAT ATGGTTCTGG AAACCTCCCCA CACCTTTCC AGATAGTTCT

-6351
GAATAACCCA CCCCTTAATT TGCATGTAAT TAAAGTCGG TATAAGTACA GTTAGCCAGC AGCCCACCTGG

-6281
CTGCTACTGT GGGCTCACTG CCTATGGGTT GTCTGCTCT GCAAGGAACA GCTACCTTGC TGCCACTGCT

-6211
GCTCAATAA ACCTGCTTTC TTCCACCACA GGCTCGCTCT TCAGTTCTTT CCTGAGCAA GTTAAGAAC

-6141
CTCCCGGGCT AAGCCCCAAT TTTGGAGCTT GCCTGCCCTG CATCAGTAGA ATGGGCTAAC TACTTACGGT

-6071

GCACTCAGGC TAAAGAGGCT GATGCTTGCA GGGCAGTATT CACAGAGCAC ACGGTAGTTC ACGGGATGCC

-6001
TCTCACCCCTT GACTCAGTGC TTAAGAAAAGG AGGGAAAATG GTGAACATGA TCAAATCATG GCCATTGCCT

-5931
ATTCATCTTT TCAGTGTGT ATGGAGGAAT AGGCAAGTAG GAGATTGCTT TTCACATTAA TGTCAAAGAG

-5861
AAAGATAGTT ACTTGGAACT TAAAAAAATT AATTGTGATA AAATATACAT AACATAAAAT TTACCATCTT

-5791
AACCATTTT AAGTATAGCC AATCTCAAGA GCTCTTCTA TCTTGTAAAA CTGAAACCCCT ATACCCATT

-5721
AACAACTCCC AATTCTCCCC TTTCCCTAAC TCCTGGCAAC CACAATTCTT TCTGTCTCTA TGAATTTGAC

-5651
TGCTTTGGCA TGTCATAGAA ATAGACTCAT ACAGCATTG TCTTTTGCG ACTGGCATAT TTTGCTTAGC

-5581
ATAATGTCTT CAAGGTTCAC CCATGTGGTA GCATGTGTCA GAATTCCCTCT CCTTTGAAG GCTGAATAAT

-5511
ATTCCATTGT GTGTATATAC CACGTTTGT TTATCCATT GCCCATCAAT GGGCATTGCG GTTGCTTTT

-5441
TTGCCTCTCA TGAATGATGA ATATGGGCGC ACAAAATATCT CTTCAAGACC ATGCTTCAA TTCTCTGGG

-5371
TATACACCCA GAAGTGGAAT TGCTGAATCA TATGGTAATT TTTTTTTTT TTTGAGACAG AATCTTGCTC

-5301
TGTTGCCAG GCTGGAGTGC AGTGGCACAA TCAGAGCTCA CTGCAGCCTT GGTCTCTGG GCTCAAGCGA

-5231
TCCTCTTGCT TCAGCCTTCC GAGCTTCTGG GACTAAAGGT GTGTGCCATC ATGCCTGGCT AATGTTTTAA

-5161
AAACGTTGCC AGGCATGGTG GCTCGTGCTT GTAATCCTAG CACTTTGGGA AGCTGAGGCA GGTGGATCCC

-5091
CTGAGGTCAG GAGTTGAGA CCAGCCTTGC CAACATGGTG AAATCCCGCC TGTACTAAAA ATACAAAAAT

-5021
TAGCTGGGTG TGGTGGCATG TGCCTGTAGT TCCAGCTACA GGCAGGAGAA TTGCTGGAAC CTGGGTGGCA

-4951
GAGGCTGCAG TGAGCCGAGA TTGCACCACT GCACTCCAGC CTGAGTGACA GAGTGAGACT CTGTCTCAA

-4881
AAAAAAAAAA ATTTAGAGA TGGTGTCTCA CTGTGTTGCC CAGGCTGGTC TTGAACTCCT GCCCTAAAGT

-4811
GATCCTCCTG CTTCCGCCTC CCAAAGTGCT GGGATTACAG GCATTAGCCA CCATGCCTGG CCTAGCTAAA

-4741

TTGTCTTAA TGTCGCATGT CTGCAAAAAA CACATCTATA AAGCTAGAAA AGTTGAGCAT CCAACTTTT

-4671

ATGATTTAAC TCTCATGACC TGGCAATT TCTAGCAAGG AGCCTGGGCT GGTGGTTTA GGAGAACTGA

-4601

GTGAAAAAAA GAAATACATT AACTAGATTG GATGCAAAGT GCCTGCTGGT CATGGGTGTT TTCTGCTGGC

-4531

CCCTGTTCAT CTGTGCCGT TAGCCCACCC ATGGGTGAGT GGGGCAAAGT GGCCAAACTG ATTCTTAAGA

N 489

-4461

GAGGCATACA TGCAGAATCC AAGTTAGTCA TGATTCGTT TCTAGTCTGA GTGAATGTGT GTCCAGAATA

-4494

-4391

TTTTATAAAC TTTATCAGCT CAGAGGGAA AACCTGTCTC CATACTACGT GGTTTATACA AAGCTGTCAG

-4321

GAATTCAAGCA TGATGAAGAA ATGCACAAAA CAAGTGTGAA CAGATAAGTA AAAGGATCTA CTGAAAATCT

-4251

TCAGGGTAGT ATATTGTGT ACAGGACCAA GAATTGAAG TCAACATCTG TATTTGTGCC CTCTGGACAA

-4181

AGGTATTATC CCTGATGATA TAAAAATTAA TTTGGGCTG GGTGTGGTGG CTCATGCCTG TAATCCCAGC

-4111

ACTTTGGGAG GCTGAGGAGG GTGAATCGAC TGACGTCAGG AGTTGGAGGC CAGCCTGTAT CGACTAATAA

-4041

TACAAAAAAA TTAGCTGGAC ATGGTGGCGT GCACCTGTAA TCCCAGCTAC TCAGGAGGCT GAGGTGGGAG

-3971

AATTGCTTGA ACTCGGGAGG CCGAGGTTGC AGTGAGTCGC ACCACTGCAC TCTAGCCTGG GCGACAGAGT

-3901

GAGACWCCGT CTCAAAATAA ACAAAATTAA TTTCGAGGCC AGGTGCAGTG GCTCCAGGTG CGGTGGTTCA

N 488

-3831

TACCTGTAAT TCCAGTGCTT TAGGAGGCCA GAGGATTGCT TGAACCCAAC AGTCGAGAT CAACCTGGC

-3761

AACATCAGTG AGACTCCATC TGTAGAAAAC AATCAAACAG ACAACAAACA ACAACAAAAA AACCAGAGGT

-3691

GGGAGGATCA CTTGAGGCCA GGAGTCCGAG GCTGCAGTGA GCTATGGTCA CGCCACTGCC CTCTAGCTTG

-3621

GGCAACAGTG CCAGACTCTG TCCTAACAA CAACAAACA AAAAATTAAT TCTACTTTAA CTGTCAGTT

-3551

CATGATATCC TTCTATTAAG AAAAACCTTT TCTATCTGAT GAACTATTGG CTAGGTTTC TTTCTCTCTG

-3481

CTTTGACTA ATGCATTTAA TTACTTTCAT TTGCAAACTC TATCCTTCTC ATCAACTTTG TATTTTAGAT

-3411

GTGTCTATTG ACAGCCTGGC TTCCCTCAGC GATCATTATG ATGATCAAAG TAGATGAATA GGTAAAATTC

-3341

AATGCAAATA TTCCAGGGCA TCTAAATCCA TACCCCAAAT GGGAAAAGGG GAGAATTGGA AGCCAGCAAT

-3271

TTGAACACAT TACTATGGAT GTATTTTCT CATGCCGGGG AAAAAGTGAT TTGGAGAGAG AGAATTATGA

-3201

ATGCATGTGA AGAATAAACGC CAAATTCCT GGGAGGAGGG GAAGACCAGG AGAAACAAAA CCAAATCCTG

-3131

GCTGTGGCCT CTAAGGCATG GGGACCTGGA GTTATGCTCT CCAGGCAGAC ACAGCTCATT CTGGAGAAAG

-3061

GCTGCAAAAAA TATTCTCCTT CACATTGATT TGAAAACAAT TATTAAATTC TTGTTTCTT ATTTATCTAA

-3013 N487

-2391

GTTGTAACCTT TTAAAACCTTA CTGAGAGAAG ACGGGCACGG TGGCTCACTG CTGTAATCCA GTACTTTGGG

-2921

AAGTCAAGGC AGGTGGATCA CCTATGGTCA GGAGTCGAG ACCAGCCTGG CCAATATGGC AAAACCCCCG

-2851

CTCTACTAAA AATACAAAAAA TTATCAGGTG TGGTGGTGTG TGCCTGTAAT CCCAGCTACT CGGGGAGGCTC

-2781

AGACAGGAGA ATCACTTGAA CCTGGGAGGC AGAGGTTGCA ATGAGCTGAG ATTGCACCCAC TGCACCTCCAG

-2711

CCTGGGCGAC AGAGCAGGAC TCCATCTCAA AATAAAAATT ATTGATTAA TTAATTAAAA ATTTACTGAG

-2641

AGCTGGTGGT TCCTTAAGG GTGGAGCCGC CATCAAGTCC CCAGAGGATG CCCTGAATTG GGGGGCATCA

-2571

CCTTCAGCTG CTGTGGACTC TGAGCCTTGG CAGCTCCAGC TCCAGGCCTG GGAGAAAGAT GATTTCTGG

-2501

CAGCGTGCAG TGATTGTGAG CATTGACTA CCTTACTGCA TTTGCCCTT ATCAKTGCTC TCCAAACATG

-2431

AGTGGAAAAC AAAAAATTG GCTGAGACAA GCGATAATAC GAGTTAGGGA AAGTTGGAGA ATTTTATAGT

-2361

TGCTGATATC AGCAAATCGT GAGTTCAAG CACTAATTAA CAGAAGGAAG TCCAAATTA AAGGGGATAT

-2291

AGAAATGTGT AAAAGATGAG GTGTGGTGA GATGGAGAAA ATGAAGAGCT CTTAAATTT CTGAATTATG

-2221

AAGAATCACC AACAAATTAT TTTGTGGTTC CAAATACAGG GAGAAGTTCA CAGATCCACA GAACTGATGA

-2151

CAGGGTGCAG CCAGCCACAA ACCTTCAGC ACAAGAGGG AAGGCTGCC GCTCCACTT GCCTGGGCAG

-2081

TCTTGTAAG GCAGTAGATA AGTCAGCCTC GAAGTTAGCA ATCACAGCCC TCGGCTCGGT TTCCTGCAAG
-282 N486 → -2011

GGCATCGTTA ATGCATCACA ATTAATTCT TCTGTCCATT AAATGTCAGC TCTCAAGTAA ATTGATGTTAA
-1941

AATTTTGTA TAGAAAACCA TTTCATATTA TTTGCACATTG ATGTTTAATT ACATTTAAA TGTTTGTT
-1871

GTTTCATTTC GTTTTGTTTG TGAGACAGAG TCTTGCTCTG TTGCCAACGC TGGAATGCAG TGGTGTGATC
-1801

TTGACTCACT GCAACCTCTG CCTCCTGGGT TTAAGCGATT CTCCGCCTC AGCTTCCTGA GTAGCTGGAA
-1731

TTACAGGCCT GCACCACCAT GCCTGGCTAA TCTTTGTATT TTTAGTAGAG ATGGGGTTTC ACCATGTTGG
-1661

CCAGGCTGGT CCCGAACCTCC TGACCTCAAG CTATACACYT GCCTCAGCCT CCCAAAGTGC TGGAATTACA
-1591

GACATAAGCC ACTCTGCCA GCCAAATGTT TTAAATAATT GTCACATATA TATACAAAAT AATTTATGTT
-1521

ATAGGTAGGG ATCTTGTAT ATTTAACCT TCAAAGTATA TTCCTAAGCT TTTTATTAT TTTTATTAT
-1451

TTATTTATTG AGACAGTCTT GCTCTGTCGC CCAGGCTGGA GTGCAGTGGC GCAATCTCGA CTCACTGCAA
-1381

ACTCTACCTC CTGGGTTCAA GCGATTCTCC TGCCTCAGCC TCCTGAGTAG CTGGGATTAC AGGTGCGCAC
-1311

CACCATGCCA AGCTAATTT TGTATTTTA GTAGAGACGG GGTTTCACCA TATTGCCAG AGCTGGTCTC
-1241

AAACTCCTGA CCTCAGGTGA TCCATCCACC TCAGCCTCTC AAAGTGCTGG GATTATAGGT GTGAGCCACT
-1171

GCGCCTGGCC TATTCCCTAGC CTTTTATATA TAGACCTTT TCTTTTCAC ATTTAAAGG AACTTTATG
-1101

TTTAATCATG GAATATTCA AACATACAGA AAAATCACAG AAAATAAATA ACAACCACTC ATTTATCTTC
-1031

TCCCCAACCC CATGTAATAA ATATTAAAAT ATTGTGTTAA ATGCTAAATT TAACACATGC TAAAGGTTCC
-961

TGGCTGGATG TGGTGGCTCA CGCCTGTAAT CCCAGTACTT TGGGAGGAGG AGGTGGGAGG ATTGCTTGAG
N485 → -891

TCCAGGAGCT CGAGACCAAGC ATGGGCAACA TAGTGCATC TCGTCTCTAC AAAAAACAAA AAAATTAGCT
-821

GGGCATGGTG GTGTGCATCA GTAATCCCAG TGACTGGAG GCTGAGGTGG GAGAATTGCT TGAGTCTGGG
-751

AATTTGAGGC TGCAGTGAGC CCTGATCATG CCACTGCATT CCAGCATGGG CGACATAGCA AAACTTGTCA

-681

AAAAAAAAAA AAGTTTCCTC TCTGCCAAC CATAGACAAC CACTCTTCTG ATTTCTATCT TCGTAGATGA

API

-611

ATTTTGCCA TTCTCTTGT A TATGAAAGGA ACCAGACATT AGGCATTCTG GTGTCTGGTT TCTTTCACCT

-541

AAGATAAAAAT TGAGTTAAC C TGTATTGTTG TACAGAACTG CAGTTGTTC TTTGTTATTT ATTGTAAAGA

-471

CAGGGTCTGG CTATGTTGCC TAGGCTGGTC TCGAACTGTT GCCCTCAAGC AATCCACCTG CCAAGCTCTG

-401

GGACCACAGG CATGAGCCAT GGCATCTGAT CKGTAGTTG ATCTTATTTC TTGCTGAGTA GTAGCCCCATG

API

-331

GCATGACTTT ATTATTTGG GTGTCCATT C TCCTCTGGAG GGGCTCTGCT TTTTGAACAC ACACCCCTGGC

Cts

-261

CTAGCTCCCC TTCTCCCTGC CTCTCTGCAG GCTCACATCC ACATGCCAAG ACCTCTGCAG CCATTCTGCT

Cts

-191

TCCTGTCCCTT CCACTCCTGT GGGACCTCAG AGAGCTACGG GGCTCCCTGG GTACCAACTG GCTCCTGAGG

S_p1 | *S_p3*

S_p1 / *S_p3*

-121

CCTGGGGGAG GGTGGTCTTC TGGGAGAAGG AAGCCAGGTC CCTGCAGGTT GTGGAGGGGG ACAGAATGAG

S_p1 / *S_p3*

Cts

-51

GGTTTTCCC CAGGATGTTG TTGGCCCCCTG CCCCCACTTC TGTTCCATAA TTAACCACGC CCCTCCTACC

S_p1 / *S_p3*

+1

20

CACTGTGCC C TCTTCTGC TGTGTGGAGG CCCTGAATCA TTATTTAAC TACCCCTGG GAGGGTGAGC

Cts

Cts

90

ACCTTCTGTG CTCTGTCCCC AACCTTCCAC TTCCCTCAA CGCGCTGCTC AGGGATGACC T TCGGCACTF

M T F G T

160

TGCTTCTCT GAGTGGTAAG TGGGGCCAGG GTGCTGGGA GAAGCTTGGGA GGAGTTCTGA GGGGACTCCA
V L L L S V

230

TCTGGGAGGG CAGGCTGGGG GCTGGTGGTC GGCTCCAACC ACTCTTATGA GGAGCTGAGG CAGGGAGTG

300

CTTCATGTGC GAGTGGCCCG GAGTCAGTAG AGTGTGACCT GAATGAAGAG GGGCTCAGGG GCTGTGCTCA

370

GGTGGGCACT AAGCTACCTC TCCAGCTGGC TATGTTGTCC CAGGCTTCCC TGCTCCACT CATGGAGTCC

440

CTGGTGTGGG TGACAGAGGT CTCCCCAGCC TCCCCCGGGA GTGGAAGGCC ACAGAAGCCA CCAGGGAGGG

510

GGAAAGGTTG GACATCACCT CCCTGGGCCT NNNNNNTCCC CCAAGTCCTG ACTGCACGTA GGGAAAGAGGC

580

CCCCCTGCTGA AAAACTGCATC AGAGTCACAT TCACGTGCCA TCAAAAATCA GGCTTGGCTG GGTGCGGTGG

650

CTCATGCTTA TAATCCCAGC ACTTTGGGAG GCCGAGATGG GCGTATCCCC TGAGGTCAGG AGTTTGTGAC

720

CAGCCTGGCC AACATGGTGA AACCCCACATCT TTACCAAAAA TATAAAAATT AGCCGGGCAT GGTGGCGTGC

790

ACTTGTAATC CCAGCTACTT GGGAGCTGA GGCAAGAGAA TCGCTTGAAC CCAGGAGACG GAAGTTGCAG

860

TGAGCTGAGA TCGTGCCGTT GCACTCCAGC CTCAGCAACA GAGCGAGACT CCATCTAAA AAAAAAAA

930

AAAAAAAGAA AAAAAAGAAA AAGAGGCTGG GAGGTCTAG GGATTGGGGC TTCTTTAACT CCCAGCCTCC

1000

CCGCCCACCA AATATTCCCTC AGTCTGGCT TCTTATCATG GATTCAACCT GGATGTGGAG GAGCCACCA
LASYHGFNL DVSEEP

1070

TCTTCAAGGA GATGCAAGC GGTCTGGCT AGAGGCTGGT GCAGTTGGT GGATCTGGT AGGCCCCACT

Inton 2
(3019)

1140

CACCCCTCCTT CCCAACCTC CACTACATCA AGTCCTGTGG ATGGGTACAC GTGGGTTACC CGAGGGAGGT

1210

GTCCTGGAGG AAGGCCAGCA GGGGTGAGAA GTCTCCCTT GGCTCCTTGG AGGCCCTGAC ATCAGCACCT

1280

ATTATTCTCA ATCCCAGGAA AGGCCACAAA ACTCTACACA AGACCCTACC TTACCTCGGG AGGGAAGCCT

1350

TGAACCTGCC TCCCAGGCAG GGCCCACCTTC TTGGGCCAG TATGGTCACA CAGGGCCAC ACTCATTAAAC

1420

TTTGGAGTTT AATGTTCTGC CCTTGACCTC TTGAAATTCC TGATTATTT TATTTTATT TTTACTCCAG

1490

CTCTGTTACC CAGGCTGGAG TGCAGTGGTG CAATCACAGC TTACTGCAGC CTCAAACCTCT CGGGCACAAG

1560

TGATCCTCTC ACCTCAGCCT CCTGAATAGC TGGGACCACA GGTGCATGCC ATCATGCCTG TTTTTGTTT

1630

TGTTTGTTT TACTTTTAC AGAGATGGAG TCTTGCTATG TTGTCCAGAC TGGCTGAACCT CCTGGGCTCA

1700

AGCAATCCTC CTGCCTTGGC CTCCCAAAGT GCTGGGATTA CAGGTGTGAG CCACCCCTGTC TTGCCAATTC

1770

TTAAAAATT TATCTGTGCA TTTGTGTTT GCAAGTAAAG AATGATGGCA GGGCTGGCA CCATGGCTCA

1840

CGCCTATAAT CCCAACSCTT TGGGAGGCTG AGGCAGGCAG ATCATCTGAG GCCAGGAGTT TGAGACCAGT

1910

TTGGCCAACA CAGCAAAACC CCATCTCTAC TAAAATGCA AAAAAAATTA GCCGGGCATG GTGGCAGGCA
1980
TCTGTAATCC CAGCTACTTG GGAGGCTGAG GCAGGAGAAT CGCTTGAACC TGGGAGGTGG AGGTTGCAGT
2050
GAGCCGAGAT CGTGCCACTT TACTCCAGCC TAGGTGACAG AGTGAGACTC CGTCAAAAAAA AAAAAGTCAT
2120
GGGAGAAGGG AGATGTACTG GGGGTTTGGA GCCTTAGCTC AGCAGCAGCC CCACCTCCCCA CCGCCTCCTG
2190
AAGGGTGGTG AAGGGGTATC AGCTGCTGGC TCCCCCACCC ATGTGGGAGC AATGACCGCT GCTACCTTCC
2260
GCCCTGGCA TGAGCTGGT AAAGTCAGTT AGGGGCGCTC ACTCTGGGAG TACCCCGAGG GAGTGGGACA
2330
CTACATAGCA AATAAAAAAC GTCAGGACAG GTTGAGGAAA GAGAGCAGAA GAAAGGTAAG AGCCCCCAA
2400
CCCCAAGAGA CCCCACAGTT TTATTTCAA TTGGGACCCA CAAATTATGA ACCTGCCCCC ACTTCCAGGA
2470
GCTCACATTG TCCTGTCCCA GAGAGTTCAA GTCACAATGT GACACAGGTG TCACCAAGGT CTGGGGGGCG
2540
CAGGCAGGGGA GAGAGCAGAC CCAGGAGGGT TCCATGGAGG AAGTGGTGCT GGCAGTGAGC CCCAGTGGAC
2610
AGGAAGGCTC AGTTGGTCAC GAGGAGCTAT AAGAGGTCAC CGAGCTCCAA CGCGCACCC CTCTCCCTTC
2680
CTCATGTGAC TGGCAGTCTG GGGGGATGGA AGCAAGCACC AGGCACCAGG CTTTGTTTT TCTTTATTG
2750
GAAATGTGGT CAACTGAGGT GCACAAATCT GAAAGACCCA ATCTGATAAA GGATACACAT GTGCGTGCCT
2820
GGGTGAGCCC CACCTAGTC AGCTGCTCCA GTGTCAAATC CCACAGGCAC AGGGCTGCCG TGGACCCCTT
2890
CTCATCACCC AACATCCCCA GAGAACCCCT GGTCAGACTT CTGTCACCAT CAGTTTTTG GGCCACATTT
2960
TAAAAAAAGA ATACATTGGC TGAGTGCAGT GGCTTATGCC TATAATCCTA GAACTTGGG AGGCTGAGGC
3030
GGGTGGATCA CCTAAGGTCA AGAGTTCAAG ACCAGCCTGA CCAATATGGT GAAACCCGT CTCTACTAAA
3100
AAATACAAAA ATTAGCCTGG CGTGATGCC GGTGCCTGTA ATCCCAGCTA GCTGGGTGAC TGAAATAGGA
3170
GATTTGCTTG AACCTGGGAG GTGGAGGTTG CAGTGAGCTG AGATCACGCC ATTGCACTCC AACCTGGGTG
3240

ACAGAGTGAA ACTCTGTCTC AAAAAACATA TGGGTTGATG GGTTACACTA AAGTTTGCT CATCGTTGT

3310

ATCAGCAGGT TCCAAACTGC TACCTCTCTA GCCAATGCTC AGATTTCTT CACAAAGCCT TAGGCATCCC

3380

CTGAATCATG ATGCACAGGG ATTGTAGCTT TCTGTAAAGG AGCGGCACCT AGAAGGAACC CTCACATGGC

3450

CATTTAATGA AGCCTTGCTT GGCGCATTAA AATACACCAG TATCTGTCTG CTCTTCAC AGACAGGAGA

3520

TTGTGGTAG TGAGAAAACA TTTCCAAAAT TAAAAAACTT TCCCACTCAG GGAGTTTGCA AAATAAACCC

3590

TTGACTCTAC ATAACATATAG ATATAGTTAT GGATCCTAGT ACAC TGCTT ACATTGCCA ATTGAAATTG

3660

CTTATACAAT ATTTAAATTG GTCCAATGAA TTACAGAATC AACTATTTGT TTTGAAAGCA CATGTCTCA

3730

GGAAATTGTT CCAATTAAC TGAGATGATC TTATTCTTG GGTGGTCAA AATAATGGCA ACTCAGAAC

3800

GCAATGTGCT TACCCATGAT TGGGAAATGC CATTGTC TTTAAATAGG TCTTTTTTT TTTTTTTTT

3870

TTTTTTTTTT GGTGAATGTT AAAAGAAAT TTCTAACAT AAATACACAC ATACGTACTT ATGCACACTC

3940

AAAACCAAAT AAACCCCAGC ATGGCCCTG GGCATCTGTG AGTTACACTT GGGCCCTGAT TTCTGAATAT

4010

TCTGCCAAGT GGCAAATGCC AGGAATTCC CCCACAGAGT CTCGCTTCCC CATGGAGGGA CACTTCCTCA

4080

CCCCCAAGTG CCCGCTGCTC CCACCCCTCC TGTGGCTGCA GTGACATGGC CATGGTTGTG TCTCCACTL

4150

cTGTTGGGA|GCACCCCTG| AGGTGGTGC|GGCCAACAG|ACGGGACGGC|TGATGACTG|CGCAGCTGCC|
V V G A P L E V V A A N Q T G R L Y D C A A A A

4220

ACCGGGATGT GCGAGCCAT CGCGCTGCAC|AGTGAGTGAC CACCTGGAA TTGGGCCCT CAACCCCTCCT

+ 6 m C Q P I P L H I

4290

Intron 3
GGACCCAAGT GTGCCCGCG TTAGCTTCCA GTCCAGACCT TCCCCGAAA TGAGTGTGTG CTGTGAGTGA

4360

GACCCCGCGT GTCTGCCCTT GCAGTCCGC CTGAGCCGT AACATGTCC TTGGCCTGA CCCTGGCACG
R P E A V N M S L G L T L A A

4430

cTCCACCAAC|GGCTCCCGC|TGCTGGAG TGAGTGTCTT GGGCCACGGG GGGGTGGGGT GGGGGCGGGGG
S T N G S R L L

4500

GTGTTGTTGG GGAGGAGGCT GGGGCTGGGA GTGAAGGAGG AGGGGCTGCT AGGGACTCCT GGCTCACAGG

Intron 4

4570

CTTCTGCCTC CAGGCGCTG GCGCGACGCT GCAAGAGTC|TGIGGGAGA ACTCATAGTC AAAGGGTCC|
A C G P T L H R V C G E N S Y S K G S
4640 (170)
TGCCTGCTGC TGEGGCTCGCG CTGGGAGATC|ATCAGACAG TCCCCGACCC CACGCCAGGT AGGTCCCTGG
C L L L 6 S R W E I I Q T V P D A T P C
4710
CAGGCCATGG TTCCCTGTGG ACCACATGCT GGCACTGAGG GTGAGCAGGC GTGAGGCCTG TGTCTGGGGC (4267)
4780
CCTGTGCCCT CCCTGGAGGG CCGAGTGTGG CTAGGAGAGA AGCCAGGAGA AGAGGGTGGC TCAGGGCAGGA
4850
GCCCTGCTGC TCCAGGGTAG AAGTTCTTG CAGGGTTTT CTTTATATT TTTCTTTT AAGACAGGGT
4920
CCCTGCCAGG CACAGTGGCT CAGGCCTGTA ATTCCAGCAT TTTAGGAGGC TGAGGTGGC GGGATCACCT
4990
GAGGTCAAGGA GTTCGAGACC AGCCTGGCCA ATGTGGTAA ACCCCTCTAC TAAAAATACA AAACAAAACA
5060
AAACAAAATA GCAGGATGTG GTGGTGTGCG CCTGTAATCC CAGCCACTCG GGTAGGCAGA GACAGAAGAA
5130
TCGCTTGAAC CCAGGAGGCG GAGGTTGCAG TGAGCTGAGA TTGTGCCATT GCACTCCAGC CTGGGTGACA
5200
AGAGCAAAAC TCCATCTCAA AAAAAAAA AAAACAAAAA ACAGAGTTTC TGTCAGGCTG CATGCACCAC
5270
CACACCCCTGC TAATTTTTT GAGACAGAGT CTTGCTCTGT CGCCCAGGCT GGAGTGCAGT GGTGCAATCA
5340
TAGCTCACTG CAGCCTCGAA CTCCCTGGCT CAAGTGATCC TCCTCCCTTA GCCTACTGAG TAGTTGGGAC
5410
TGCAGGTACA TGCATCACAC CTGGCTAATT AAAAAAAATG TTTTGTAGA AATGGGGTC TTGCTATGTT
5480
ACCCAGCCTG GTCTTGAAC CCTGGGCTCA AGTAATCCTC TGCCACAGGC TCTCAAAGTG TTGGGATGAC
5550
AGGCATGAGT CCTTGTGCCT GCCCTGAGGG ATGAAAGTTC TGATGGAGGC AGAGAGGAGC CCCACTGTGC
5620
GGGCTGTAGA GGGCACAGCA TCTTCCAGTT GCCAACAGGT GCATGGCCAC TTCTTGAGTT TCAGAGGAAG
5690
GACCTTAGTG TGGTAAAGAA CGTGGTGAGG AAGATAAACATC CATGAGGGAG GTGTTCTTC TGGATGGTTC
5760
ACTGCTGAGC TTCCAGGATT CCCCAAACCA ACTTTCTCT CGAAGAGGAG CAAATGACAG GGCTGCGGAA
5830
AATGCGATGT GCAATTTGT CAGTGCCAT GTCTTCCACA GAGAACAGGG CCTGGGGACA CCACCATGAC
5900

ATCTCTCTGA GGGTTGGTCT GCATCATGGT GGTTCCAAG TTTGTTTCC ATGGGCACCA GGCTTCATT
5970
CCTTGAAGCT TCATTCCCTC AAAGCCATTG AGTTCCCTCA TTGGTAAAAT AGAGCTCAAT AATCAGGGGG
6040
TTATGAAGGT GAAAGGGATT GAGGTGCATA AAGCACTTGG AACCCCTGCCT GGCACATAGT ATGTGATAGC
6110
CCCTCTGACC CATCTTCCAG CTGGGGACTG CATGCTGGGA CTGGGAGGAA GATACAGGCA AACTGTCTCA
6180
TCTGCCGTGT GAGAGGAAT GCCAGGGGCC GCTCAGGGTG CTGACCGAGG GTGGGGCTTC AGACCAGAGA
6250
GGCCATGATG ACAGGCATGC TGGGCCTTA GACAAAGGTG GAGCAGCAGC AGAACACATTA CCAGAGCAA
6320
TGGTGAGGGT GGAGTCTATG GAGGGGACCA AGGCAAGGGG GAAGGGACAT CCAGGGTTCT TGGGGGGACC
6390
GTGCCAGCC TGAGATGTCT GTGAAGCTAG GTTAGGGAGG TGGCACTTAA AAACAAGGGG TAAATGTCTT
6460
CTCACAGCCA TCCGTGGAAC TCATGAGGTG GGATGCCTGA TGCAAATGGG ACTGGAGCAC AAAACTGGTG
6530
CAGGCAAGGG GGGTGTGGGT CCAAGTAGAA GGGACCAGGG TCCACTGAGG ATCACCTGTG TGCCAAGCAG
6600
TGCTGAATAC CTGGTATGAA TCACCTTATT GCATCCTCAC AACATCCTGG GTGGTGGCA GGCCCATTCT
6670
CATTTCACAG ATATGAAAAC CAAGGTTCAAG ATAGATGAGT TCCATCGATA GCAAGAGGCA GAGCCCAGAG
6740
CTTGAGCCAT CCTTGCCCTGA TTGGTGGGT CCTTTTCAA AAGGATAAGT CCAGGCTTCT GCTAGTGGGA
6810
GACCAGGGGA TACAATAAAA AGACCAAGAA ACAGAAGAGA CATTGTGAGA GGATTGCCA CAGACCTGGC
6880
CTGAGAGAGG ATGAGAGGGT GGTTCTTGA CGCAGCTGAA AAAACAGGCA CCACTGCAAG ATGTTGGCTG
6950
CCCAGATGTG GGCAAAAAAC GGGGAGCTCC TGGGGGGATC TGCAGCCTGC CCCATGGATG TCAAGATTTG
7020
CTGGTGATTG AAGAACAGG AAGGAAGTGA CCTTCTGTTT CTCCCCAGCA CCCTTGAGC ACCAGTGGTT
7090
GAGCAAGTGG GGTAGGGGAG AGGAAAGAGG AAAAGGCATT TTTTTTTCT GCAGTGGTGG GCAGGGGGCA
7160
GAAACCACAG CCCTGTGGTG TGGGCCTCAC ACCTTAGTGC TCTGGTGGCC TGATCTCCCA GTGCCCTGCG
7230

GGCAGCACAG GATGTGGCTG CTGGTGGAGG TACCAACTGG GCCCTGAACA CAGGCCACAC ACCCCCCATG
7300
AGCCTGGGGA CAGCATGAAA AGTCTTATTGTT CATGTGC ATATGATGTG CCCTCACGAT TGCAGAGTGA
7370
ACTCCACAAA CTCTGAGGTC ACTTGGGAAT GTTCTTTTT TTTGAGACGG AGTCTCACTC TGTCGCCCAG
7440
GCTGGAGTGC AGTGGCACAA TCTTGGCTCA CTGCAGCCTC CACCTCCCAG GTTCAAGTGA TTCTCCTGCC
7510
TCAGCCCCC AAATAGCTAG GATTACAGGC ACCGCCACCA TGCCGGCTA ATTTTTTGT ATTTTTAGTA
7580
GAGATGGGGT TTCACCAGT TGGCCAGGCT GGTTTGAAC TCCTGACCTC AAGTAATCCG CCCACCTCAG
7650
CGTCCCAAAG TGCTAGGATT ACAGGCCTGA GCCACCACTC CCACACTGGGA ATATTCTTG GCACCGCACC
7720
CATGGGAGCA TGAAGGGTGG ATGCAATGCA ATCATAAACAG AGGCCAAGG TCAGCACTGG GGTGCTTGGC
7790
TGTCACTCCA GTGCTTGGG AGGCCGAGGT GAGTGGATCG TTAGAGCCCA GGAGGTTGAG ACCAGCCTGG
7860
GCAACATGGC GAAACTCCGT CTCTACAAAA AGATACAAAA ATTAGCCAGG CAAGGTGGTG CACACCTGTA
7930
GTCCCAGCTA CTCAGGAGAC TGAGGTGGGA GAATTGCCTG AGCCTGGGGA GGTCGAGGCT GCACTGACCT
8000
GTGATCACAC CACCACACTC CAGCCTGGGT GACAGTGAGA CTCTGCCTCA AAAAAACAAA AAATGAAAAA
8070
ACCAGAGGCC TCAGCCAATG CCTGGGGGCC TCAGAGTGCA GCTGGCCCTT CAGACGCTGA ACCAGTCATC
8140
GGTAAAGGTT TCCTCCAGGG GCAGGAGGTG TCCCAGTGGG CAACAGTTCC CCTCTGCCTA GCGTGATTCC
8210
TGGGAAGGGA CTCAGCTCAG AGCCAATC ACAGTAGCTGG AAATAAGGAC CTCTGACCGA CTGGGGGTAG
8280
GGTGGGGTCT GGGGTGGATC CCTGCCAAC CCCCACAGCA TCCCTACAGG CATATCCTAC AGGCCTCGAA
8350
GGTGCCTGGC ACAGTGGTGA AATGGTGCCA GCGGCTGACC CTGGCAGAGG GCCAGGACTT GTCTCCAGCA
8420
CCCCATGTGCG TGTTGCCTTA TCCTTGAGT GATCCCACGT GGTAGCCACT GATATTACCT TCATTTTACA
8490
GATAGGGACA CTGAAGTCCA GAGAAGTTAA GTAATGTGCC TGAATTCAAC AATTAGCACG TGGCTGAGCT
8560

GGGGTTTAGC CGGGCACAC TGGCTCCAGA ACACGGGCC TGAAACACTT TGCTAACAT TCGCCCTTGA

8630
TGCTTGTGGC ACCCCTAGCA TCTGTGTTA ATGAATATTG TTGATTAAA TGGATGAGGA GCCCCACCTGG

8700
GTCCTGTGTT GTCATCCCTC TCTTCAGG CCATGTGGGA GGAAGGGAGC AGGGGGCTGG GGTGGCAGAC

8770
TGGGGCCTCC TCCAAGGAGG GGTGGAAAC TAGGTGGGA TGCCAAGAAC AGCCCCGGG CTCTGTTGAG

8840 (1773)
CAGGAGCTGC AGGAGGGGT TGGGCCCCCG CAGTGCATCT CCGATTCCCTC CCCATTCCCC CACAGACTG|

8910
CCACATCAAG AGATGGACAT CTGCTTCTG ATTGAAGGAT CTGGAGGAT TAACTAAAT GACTTTAACG
P H Q S M D I V F L I D G S G S I D Q N D F N Q

8980 (1714)
AGATGAAGG CTTTGTCCTAA GCTGTCATG| GCCAGTTGA GGGACTGAC| ACCCTGGTGA AGACTGGCA
M K G F V Q A V M 6 Q F E G T D T L Intron 1 (1255)

9050
CCTGGGCTG CGGTTGGGG GACGGGGAG GCTGGCCTCG GGGAGGCATC CCGGGAGGGG TGGGGCAGG

9120
CCAGTGAGCC GTGTGTGATG GGGCTGGGT GGAGAATGAA GCTATGGTCC CAGCACAGGC CCAACTTGAG

9190
CCCTGACCTA GGAGGCCCTC TTGGCATTAA ATAATGTATC CAAAGCTAC AGGAAATACA ATGTTCTGC

9260
TGTAAAGG CATTGAAAGT GTTAAAGG TAATTTGCT ATGGAATTG ACTGACATAT TTTCTATGTA

9330
ACTAATGATT TCTTTTATT ATTTTATTTC CCCATAACTA GTGCCTATGT ATGTGAAAA AAACCCCACT

9400
GTTTGTGTT TTAAACAGGG TCTTGTCTG TCATCCAGGC TGCAGTGTGG GGGTGCATC ATAGCTCACT

9470
GCAGACTTGA TCTCCCAGGC TCAAGTGATC CTCTTCCTC AGCCTCCTGA GTAGATGGGA CTACAGGTGC

9540
ATGCCACCAT GCCTGGCTAA TTTTAAAGG TTTTATTGT AGAGATGGGG TCTTGTATA TTGCCAAC

9610
TGGTCTCCAA CTCCTGGCTC AAACAACCT CCCACCTTGG CCTCTCAAAG TGCTGGGATT ACAGGCATGA

9680
GTCACTACGC CCAGGCCACAA ATTCTGAGA AGGGAGGAAA TACACAAATT GATTTAATGT GTGATAAAAT

9750
TTTAAAGG CTAAGGGTAG CAAATATTGT GTTGGCAA GTTATTTTA ATAAACTGTG AAGGAAATAT

9820
CATTTGGTT AGATTTCCA GTTATTTTGT GTTGGCAA GTTGGTTCTC TGGGTTTCAG

9890

TATATTATTT GTAGGGCTTC AAAACACATG GGAAATGGTT TGTAAATCCA AAATAATTCC AAAATAAAGT

9960
TTATTAAAAC TGAAAACAAT ATGGCTTGGT GTGGTGGCTC ACACCTGTAA TCCCAGCACT TTGGGAGGCT

10030
NGAGGTGGGA GTATTGCTTG AGGCCAAGAG TTCGAGACCA GCCTGGCAA CATA GTGAGA CCTTGTCTCT

10100
ACCAAAAAACA AAACAAAACA AAAAACAAAG CCAGGCATGT GACGTGTGCC TGTAGTTCCA GCTACTTGGA

10170
GGCTGAGGCA GGAGGATCAC TTGAGGCCAG GAGTTGAGA GACCCTGTCT CTACAAAAAA TTAAAATAAA

10240 (144)
AACAAATAGTA ACAGGCAC TG AGCCCTGGC CCTCCCCACT GCCCTTGCA ~~GTTTGCAC~~ ~~ATG~~ CAGTA
~~F A L M Q Y~~

10310
S N L L K I H F T F T Q F R T S P S Q Q S L V D

10380
P I V Q L K G L T F ~~A~~ ~~H~~ ~~Z~~ I ~~Z~~ T V

GACCCCA... Intron 7

T G L